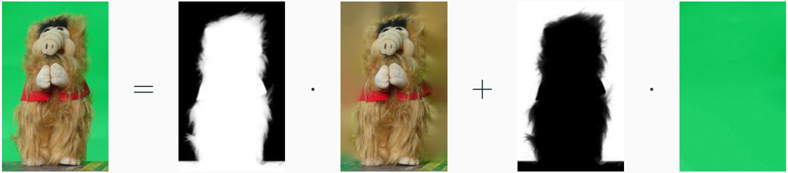
# Fast Multi-Level Foreground Estimation

Thomas Germer, Tobias Uelwer, Stefan Conrad, Stefan Harmeling

Heinrich Heine
University Düsseldorf

# Image Composition

$$I_i \quad = \quad \alpha_i \quad \cdot \quad F_i \quad + \quad (1 - \alpha_i) \quad \cdot \quad B_i$$



- $I_i \in \mathbb{R}^3$: Image color (at pixel position $i$)
- $\alpha_i \in \mathbb{R}$ : Alpha value (translucency)
- $F_i \in \mathbb{R}^3$: Foreground color
- $B_i \in \mathbb{R}^3$: Background color

# Foreground Estimation

$I$

and

$\alpha$

$\rightarrow$

$F$

- Goal: Obtain foreground $F$ from image $I$ and alpha matte $\alpha$
- Problem: Underconstrained
  - 6 unknowns in $F_i$ and $B_i$
  - 3 equations (one for each color channel)

# Motivation

$\alpha I$

vs

$\alpha F$

- Naively composing image $I$ onto background leads to color bleeding
  - $\alpha I = \alpha(\alpha F + (1-\alpha)B) = \alpha^2 F + \alpha(1-\alpha)B \neq \alpha F$
- Foreground $F$ required to achieve pleasant results

# Closed Form Foreground Estimation [LLW07]

$$\underset{\text{global}}{\text{cost}}(F, B) = \sum_{i \in I} \sum_{c} \left[ \alpha_i F_i^c + (1 - \alpha_i) B_i^c - I_i^c \right]^2 + |\alpha_{i_x}| \left[ \left( F_{i_x}^c \right)^2 + \left( B_{i_x}^c \right)^2 \right] + |\alpha_{i_y}| \left[ \left( F_{i_y}^c \right)^2 + \left( B_{i_y}^c \right)^2 \right]$$

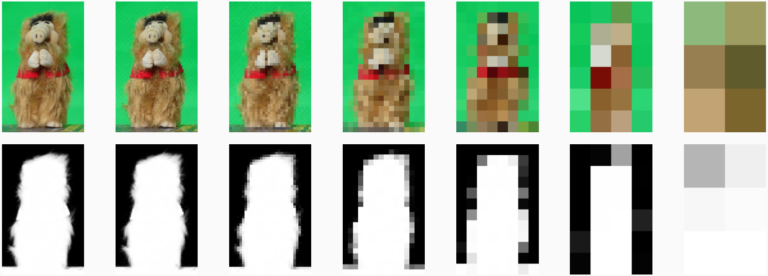
- Minimize cost function over pixels $i$ and color channels $c$
- Constrain composite color
- Penalize horizontal color gradients $F_{i_x}$, $B_{i_x}$ in regions of large alpha gradients $|\alpha_{i_x}|$
- Same for vertical color gradients $F_{i_y}$, $B_{i_y}$
- Requires 3 solves (one per color channel) of a $2N \times 2N$ sparse linear system where N is number of pixels

# Solution? - Local Formulation of Method by [LLW07]

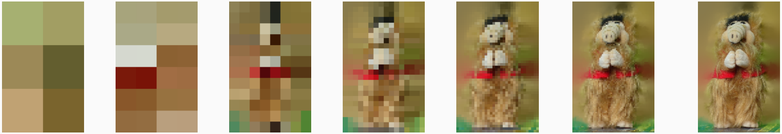
$$\underset{\text{local}}{\text{cost}}(F_i^c, B_i^c) = (\alpha_i F_i^c + (1-\alpha_i) B_i^c - I_i^c)^2$$
$$+ \sum_{j \in N_i} (\epsilon_r + \omega |\alpha_i - \alpha_j|) \left[ (F_i^c - F_j^c)^2 + (B_i^c - B_j^c)^2 \right]$$

- Local cost function over neighbors $j \in N_i$ of pixel $i$
- Control regularization with parameter $\epsilon_r$
- Weight gradient term with parameter $\omega$
- Problem: How to solve this cost function?
  - Iterative approach infeasible
  - Solution only propagates slowly across image

# Solution - Multi-Level Approach

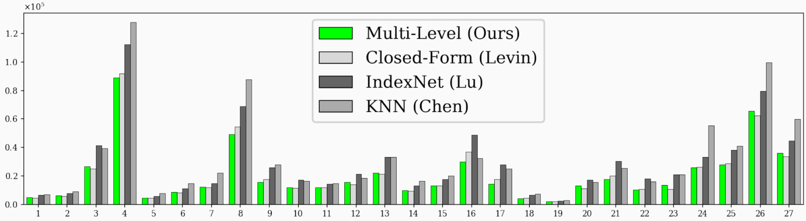


1. Downsample input image and $\alpha$ until small
2. Solve at lowest resolution, use as initialization for larger size



2x3 4x6 8x12 16x25 32x51 64x102 128x204

# Quality of Estimated Foreground




- Computed sum of absolute differences (SAD) for 27 images in dataset by [RRW+09]
- Adapted IndexNet by [LDSX19] to perform foreground estimation instead of alpha matting

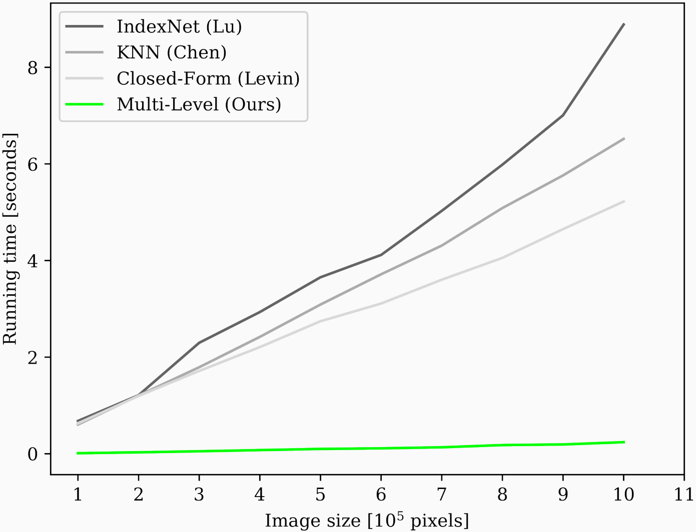
# Quality of Estimated Foreground for Various Alpha Mattes

| Alpha | Foreground | SAD $10^{-3}$ | MSE $10^{3}$ | GRAD $10^{-3}$ |
|---|---|---|---|---|
| $\alpha_{gt}$ | Multi-Level (Ours) | **20.9** | 1.44 | 8.89 |
| | Closed-Form (Levin) | 21.1 | **1.34** | **8.13** |
| | IndexNet (Lu) | 28.8 | 2.33 | 11.1 |
| | KNN (Chen) | 32.0 | 3.25 | 16.1 |
| $\alpha_{KNN}$ | Multi-Level (Ours) | **31.8** | **2.5** | **11.5** |
| | Closed-Form (Levin) | 36.6 | 3.51 | 14.2 |
| | IndexNet (Lu) | 38.3 | 3.9 | 14.5 |
| | KNN (Chen) | 34.6 | 3.22 | 13.0 |
| $\alpha_{IDX}$ | Multi-Level (Ours) | 47.9 | 5.66 | **15.8** |
| | Closed-Form (Levin) | 59.0 | 8.03 | 21.5 |
| | IndexNet (Lu) | 62.6 | 8.65 | 21.4 |
| | KNN (Chen) | **37.1** | **3.81** | 16.9 |
| $\alpha_{IFM}$ | Multi-Level (Ours) | **31.6** | **2.44** | **11.4** |
| | Closed-Form (Levin) | 37.7 | 3.98 | 15.3 |
| | IndexNet (Lu) | 36.4 | 3.93 | 15.7 |
| | KNN (Chen) | 33.7 | 2.97 | 13.6 |

# Average Runtime per Image

| Setup | Method | Time [s] | Std. dev. [s] |
|---|---|---|---|
| HPC | Multi-Level (Ours) | **2.04** | 0.296 |
| | Closed-Form [LLW07] | 26.3 | 5.48 |
| | IndexNet [LDSX19] | 74.5 | 10.1 |
| | KNN [CLT13] | 38.2 | 6.47 |
| MacBook | Multi-Level (Ours) | **1.48** | 0.251 |
| | Closed-form [LLW07] | 27.9 | 7.93 |
| | IndexNet [LDSX19] | – | – |
| | KNN [CLT13] | 148.0 | 56.2 |

# Runtime for Varying Image Sizes



# Maximum Memory Usage

| Method | Memory [MB] | Data Type |
|---|---|---|
| Multi-Level (Ours) | **1 182** | 64-bit float |
| Closed-Form [LLW07] | 7 781 | 64-bit float |
| IndexNet [LDSX19] | 91 648 | 32-bit float |
| KNN [CLT13] | 7 850 | 64-bit float |

- IndexNet evaluated with 32-bit float precision due to high memory usage

# Summary & Conclusion

- Multi-level approach effective to solve local cost function for foreground estimation
- Comparable quality to existing methods
- Order of magnitude faster
- Low memory usage
- Robust with respect to different alpha estimates

# Implementation

- Part of open-source PyMatting library [GUCH20]
  `https://github.com/pymatting/pymatting`
- Also available with OpenCL or CUDA acceleration
- Easy installation via `pip install PyMatting`

```python
from pymatting import *
# Load images
image = load_image("image.png", "RGB")
alpha = load_image("alpha.png", "GRAY")
# Estimate foreground
foreground = estimate_foreground_ml(image, alpha)
# Concatenate RGB and alpha channels
foreground_with_alpha = stack_images(foreground, alpha)
# Save resulting image
save_image("result.png", foreground_with_alpha)
```

# References


Yagiz Aksoy, Tunc Ozan Aydin, and Marc Pollefeys.
**Designing effective inter-pixel information flow for natural image matting.**
In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 29–37, 2017.

Qifeng Chen, Dingzeyu Li, and Chi-Keung Tang.
**KNN matting.**
*IEEE transactions on pattern analysis and machine intelligence*, 35(9):2175–2188, 2013.

Thomas Germer, Tobias Uelwer, Stefan Conrad, and Stefan Harmeling.
**Pymatting: A python library for alpha matting.**
*Journal of Open Source Software*, 5(54):2481, 2020.

Qiqi Hou and Feng Liu.
**Context-aware image matting for simultaneous foreground and alpha estimation.**
In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4130–4139, 2019.

Hao Lu, Yutong Dai, Chunhua Shen, and Songcen Xu.
**Indices matter: Learning to index for deep image matting.**
In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3266–3275, 2019.

Anat Levin, Dani Lischinski, and Yair Weiss.
**A closed-form solution to natural image matting.**
*IEEE transactions on pattern analysis and machine intelligence*, 30(2):228–242, 2007.

Christoph Rhemann, Carsten Rother, Jue Wang, Margrit Gelautz, Pushmeet Kohli, and Pamela Rott.
**A perceptually motivated online benchmark for image matting.**
In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1826–1833. IEEE, 2009.

Jingwei Tang, Yagiz Aksoy, Cengiz Oztireli, Markus Gross, and Tunc Ozan Aydin.
**Learning-based sampling for natural image matting.**
In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3055–3063, 2019.